Best Practices for a Future Open Code Policy for NASA Space Science
Response to a call for White Papers

Dr Peter Young (pyoung9@gmu.edu)
George Mason University
NASA/Goddard Space Flight Center
Northumbria University, UK

I write this White Paper in my position as one of the five team members of the CHIANTI atomic database project (http://chiantidatabase.org). CHIANTI is a package of atomic data and software that allows scientists to compute synthetic spectra of astrophysical objects under the optically thin approximation. It has been widely used by the Heliophysics and Astrophysics communities, with over 3000 citations to our papers. In addition, it has been used by NASA instrument teams as a component for data analysis software and for calibration. The data from CHIANTI has also been incorporated into other plasma codes, such as CLOUDY, XSTAR, ATOMDB and MOCASSIN.

From the first release back in 1996, we made both data and software freely available to the community and I strongly feel that this decision helped make CHIANTI such a success. CHIANTI has been applied to scenarios that we never envisaged, and by people from all around the world, many of whom the team has not even met. If we had kept CHIANTI private and required that users collaborate directly with the team, then the usage would have been much lower and scientific advances would have been impeded.

CHIANTI is a prime example of how an open data/open software policy can benefit the scientific community, and I strongly recommend that NASA adopt an open code policy.

Disadvantages of the CHIANTI open policy

Whereas the open policy has made CHIANTI a great success and I would strongly encourage any other modelers to follow the CHIANTI team's example, I highlight potential problems that NASA should address.

If a modeler maintains proprietary rights to his/her code, then this has advantages when it comes to securing funding. If scientists want to use the code as part of a project, then they have to obtain the modeler's services as a co-I on the proposal. This funding then helps to secure the modeler's position.

With CHIANTI, a scientist has free access to CHIANTI and so does not need to include a CHIANTI team member on their proposal. CHIANTI team members thus receive no direct funding support for their work, even though CHIANTI may play a crucial role in the scientist's project.

After 20 years, CHIANTI has reached a mature stage whereby the team is mostly performing relatively modest updates such as adding improved atomic models, and optimizing software. These changes are often not sufficient to make a compelling science proposal on its own merit, yet they are critical to assuring the continued quality of the database that users value.

Unfortunately, there is no opportunity within the current Heliophysics program to propose for the maintenance of existing codes and so the CHIANTI team has not been successful in securing Heliophysics funding despite the fact that this field includes the largest number of CHIANTI users.

Solution: mission-style funding for critical codes

My suggestion is that NASA create a funding line that can assign funding to mature, science-critical modeling codes in a similar manner to that in which NASA missions are funded. That is, as long as a code is being actively used by the science community as determined through metrics and reviewed by a panel then the development team receives funding to *maintain* the code.

I suggest that this process be specifically applied to code maintenance. Significantly new features to the code that will enable new science should be proposed through standard NASA science grants.

To qualify for funding, the code team must satisfy certain criteria. For example,

- Documentation is available through the internet.
- All code and data are freely available.
- Communication channels to the code team are available (e.g., e-mails, Google groups, online forms, etc.).
- Codes and their updates are described in the refereed literature.

The codes should be reviewed every 2 to 3 years, and proposals submitted. Some criteria by which the codes should be judged include:

- Usage statistics, such as citations to the code paper(s), code download statistics, user statistics (e.g., membership of a mailing list or help group).
- Examples of key science works that have used the code, with explanations of how the code was critical to the science.

The available funding need only be fairly modest. For example, funding levels of 10-30% for the three US-based members of the CHIANTI team would probably be sufficient.

My proposed funding model should not be seen as a form of welfare support. If codes cease to be useful for the community, then funding should be removed. Equally, newer codes that transition to the mature phase should be able to access funding if they are competitive with funded codes.